

# A Game-theoretic Perspective on Communication for Omniscience

Ni Ding\*, Chung Chan<sup>†</sup>, Tie Liu<sup>‡</sup>, Rodney A. Kennedy\*, Parastoo Sadeghi\*

**Abstract**—We propose a coalition game model for the problem of communication for omniscience (CO). In this game model, the core contains all achievable rate vectors for CO with sum-rate being equal to a given value. Any rate vector in the core distributes the sum-rate among users in a way that makes all users willing to cooperate in CO. We give the necessary and sufficient condition for the core to be nonempty. Based on this condition, we derive the expression of the minimum sum-rate for CO and show that this expression is consistent with the results in multivariate mutual information (MMI) and coded cooperative data exchange (CCDE). We prove that the coalition game model is convex if the sum-rate is no less than the minimal value. In this case, the core is non-empty and a rate vector in the core that allocates the sum-rate among the users in a fair manner can be found by calculating the Shapley value.

## I. INTRODUCTION

Communication for omniscience (CO) is a problem proposed in [1]. It is assumed that there is a group of users in the system and each of them observes a component of a discrete memoryless multiple source in private. The users can exchange their information in certain way, e.g., communicating over lossless broadcast channels, so as to attain *omniscience*, the state that each user obtains the total information in the entire multiple source in the system. The CO problem in [1] is based on an asymptotic source model. The coded cooperative data exchange (CCDE) problem proposed in [2] is a special case of the CO problem where the source model is a non-asymptotic finite linear packet one. The non-asymptotic model differs from the asymptotic one in that the communication rates only take integer values. By allowing packet splitting, the CCDE problem has been extended for asymptotic model in [3], [4]. Independently, the same model has been considered in the closely related secret key agreement (SKA) problem by [5], and is called the finite linear source model.

Determining the minimum sum-rate and finding an optimal rate vector that allocates the minimum sum-rate such that omniscience is achievable are two fundamental problems in CO. The expressions of the minimum sum-rate for asymptotic and non-asymptotic models are derived in [6] based on multivariate mutual information (MMI) for SKA [7] and in [4] for CCDE. It is also shown that an optimal rate vector can be solved

via submodular function minimization (SFM) algorithms in strongly polynomial time, e.g., the algorithms in [3], [8] based on Edmond's greedy algorithm [9]. However, these works only focus on finding a solution for CO while neglecting the users' motivation to cooperate. For example, the algorithms in [3], [8] usually output an unfair rate vector which may discourage some users from taking part in CO.

In this paper, we view the CO problem from a coalitional game theoretic perspective. In this game model, each user is assumed to be self-determined in that they can decide whether or not to cooperate and join a certain coalition, a subset of the users. The core of the game is the set of achievable rate vectors with sum-rate being equal to a given value in CO and any rate vector in the core distributes the sum-rate among users in a way that makes all users willing to cooperate in the grand coalition, the entire user set. By using the concepts and related results of submodularity and its base polyhedron [10], we derive a necessary and sufficient condition for the core to be nonempty. We give the expressions of the minimum sum-rate for asymptotic and non-asymptotic models and show that they are in agreement with the results in [4], [6]. The coalitional game model also addresses another problem in CO: how to allocate the rate fairly to motivate the users to cooperate in asymptotic model. We show that the game is equivalent to a convex game and a fair rate allocation method can be determined by Shapley value if the sum-rate is no less than the minimum value. As compared to the existing method that addresses the fairness in CCDE in [11], the main advantage of Shapley value is that it can be calculated in a decentralized manner, i.e., it is possible for each user to obtain his/her tuple in Shapley value by him/herself.

## II. SYSTEM MODEL

Let  $V$  be a finite set. We assume that there are  $|V| > 1$  users in the system. A random variable will be denoted by the sans serif font as in  $Z$ , and its alphabet by the usual math font as in  $\mathcal{Z}$ .  $Z_V = (Z_i : i \in V)$  is a vector of discrete random variables indexed by  $V$ . For each  $i \in V$ , user  $i$  can privately observe an  $n$ -sequence  $Z_i^n$  of the random source  $Z_i$  that is i.i.d. generated according to the joint distribution  $P_{Z_V}$ . We allow users exchange their sources directly so as to let all users  $i \in V$  recover the source sequence  $Z_V^n$ . We consider both asymptotic and non-asymptotic models. In the asymptotic model, we will characterize the asymptotic behavior as the *block length*  $n$  goes to infinity. In non-asymptotic model, the communication rates are required to be integer valued.

Let  $r_V = (r_i : i \in V)$  be a rate vector indexed by  $V$ . We call  $r_V$  an achievable rate vector if omniscience is possible by

\*Ni Ding, Rodney A. Kennedy and Parastoo Sadeghi are with the Research School of Engineering, College of Engineering and Computer Science, the Australian National University (email: {ni.ding, rodney.kennedy, parastoo.sadeghi}@anu.edu.au).

<sup>†</sup>Chung Chan is with the Institute of Network Coding at Chinese University of Hong Kong (email: cchan@inc.cuhk.edu.hk).

<sup>‡</sup>Tie Liu is with the Department of Electrical and Computer Engineering, Texas A&M University (email: tieliu@tamu.edu).

letting users communicate with the rates designated by  $r_V$ . Let  $r(V) = \sum_{i \in V} r_i$ . For  $X, Y \subseteq V$ , let  $H(Z_X)$  be the amount of randomness in  $Z_X$  measured by Shannon entropy [12] and  $H(Z_X|Z_Y) = H(Z_{X \cup Y}) - H(Z_Y)$  be the conditional entropy of  $Z_X$  given  $Z_Y$ . It is shown in [1] that an achievable rate must satisfy the Slepian-Wolf constraints:

$$r(X) \geq H(Z_X|Z_{V \setminus X}), \quad \forall X \subset V. \quad (1)$$

The interpretation of the Slepian-Wolf constraint on  $X$  is: To achieve CO, the total amount of information sent from user set  $X$  should be at least complementary to total amount of information that is missing in user set  $V \setminus X$ . The set of all achievable rate vectors is

$$\mathcal{R}_{\text{CO}}(Z_V) = \{r_V \in \mathbb{R}_+^{|V|} : r(X) \geq H(Z_X|Z_{V \setminus X}), \forall X \subset V\}$$

and the minimum sum-rate can be determined by the following linear programming (LP)

$$R_{\text{CO}}(Z_V) = \min\{r(V) : r_V \in \mathcal{R}_{\text{CO}}(Z_V)\}.$$

We denote the set of optimal rates as

$$\mathcal{R}_{\text{CO}}^*(Z_V) = \{r_V \in \mathcal{R}_{\text{CO}}(Z_V) : r(V) = R_{\text{CO}}(Z_V)\}. \quad (2)$$

For non-asymptotic model, the achievable rate set is  $\mathcal{R}_{\text{CO}}(Z_V) \cap \mathbb{Z}^{|V|}$ ,  $R_{\text{CO}}(Z_V) \in \mathbb{N}_0$  and the optimal rate set is  $\mathcal{R}_{\text{CO}}^*(Z_V) \cap \mathbb{Z}^{|V|}$ .

In CCDE, it is assumed that user  $i$  obtains a packet set  $W_{\{i\}}$  that contains finite number of packets each of which belongs to a field  $\mathbb{F}_q$ . The users transmit linear combinations of their packet set via lossless wireless channels so as to help the others to recover all packets in  $W_V = \cup_{i \in V} W_{\{i\}}$ . In this problem, the value of entropy function  $H(Z_X)$  can be obtained by counting the number of packets in  $W_X$ , i.e.,  $H(Z_X) = |W_X|$  and  $H(Z_X|Z_Y) = |W_{X \cup Y}| - |W_Y|$ , and all results derived in this paper hold.

**Example II.1.** Consider the set  $V = \{1, 2, 3\}$  where 3 users observe respectively

$$\begin{aligned} Z_1 &= (W_a, W_b, W_c, W_d, W_e), \\ Z_2 &= (W_a, W_b, W_f), \\ Z_3 &= (W_c, W_d, W_f). \end{aligned}$$

$W_i$ 's are independent uniformly distributed random bits. The users exchange their private observations to achieve the global omniscience of  $Z_V = (W_a, \dots, W_f)$ . In this system, the achievable rate set is

$$\begin{aligned} \mathcal{R}_{\text{CO}}(Z_V) &= \{r_V \in \mathbb{R}_+^{|V|} : r(\emptyset) = 0, \\ &\quad r(\{1\}) \geq H(Z_1|Z_{\{2,3\}}) = 1, \\ &\quad r(\{2\}) \geq H(Z_2|Z_{\{1,3\}}) = 0, \\ &\quad r(\{3\}) \geq H(Z_3|Z_{\{1,2\}}) = 0, \\ &\quad r(\{1,2\}) \geq H(Z_{\{1,2\}}|Z_3) = 3, \\ &\quad r(\{1,3\}) \geq H(Z_{\{1,3\}}|Z_2) = 3, \\ &\quad r(\{2,3\}) \geq H(Z_{\{2,3\}}|Z_1) = 1\}. \end{aligned}$$

It can be shown that  $R_{\text{CO}}(Z_V) = 3.5$  and  $\mathcal{R}_{\text{CO}}^*(Z_V) = \{(2.5, 0.5, 0.5)\}$  for the asymptotic model and  $R_{\text{CO}}(Z_V) = 4$  and  $\mathcal{R}_{\text{CO}}^*(Z_V) \cap \mathbb{Z}^{|V|} = \{(3, 0, 1), (2, 1, 1), (3, 1, 0)\}$  for the non-asymptotic model.

### III. COALITIONAL GAME

We model the system as a coalitional game  $G(V, \alpha, f_\alpha)$ . In this game, it is assumed that the users can choose to cooperate and form coalitions. A coalition is a group/set of clients that is denoted by  $X \subseteq V$  and  $V$  is called the *grand coalition*. Let  $\alpha \in \mathbb{R}_+$ . We define the *characteristic function* for a given value of  $\alpha$  as

$$f_\alpha(X) = \begin{cases} H(Z_X|Z_{V \setminus X}) & X \subset V \\ \alpha & X = V \end{cases}.$$

We call  $f_\alpha(X)$  the *value* of coalition  $X$  which quantifies the payoff of forming coalition  $X$ .

#### A. Core

The *core* of  $G(V, \alpha, f_\alpha)$  is

$$\mathcal{C}_G = \{r_V \in \mathbb{R}_+^{|V|} : r(X) \geq f_\alpha(X), \forall X \subseteq V, \\ r(V) = f_\alpha(V)\} \quad (3)$$

for asymptotic model and  $\mathcal{C}_G \cap \mathbb{Z}^{|V|}$  for non-asymptotic model. Let the polyhedron of the characteristic function  $f_\alpha$  be

$$P(f_\alpha, \geq) = \{r_V \in \mathbb{R}_+^{|V|} : r(X) \geq H(Z_X|Z_{V \setminus X}), \forall X \subseteq V\}.$$

The core  $\mathcal{C}_G$  is exactly the base polyhedron of  $f_\alpha$ :

$$B(f_\alpha, \geq) = \{r_V \in \mathbb{R}_+^{|V|} : r_V \in P(f_\alpha, \geq), r(V) = f_\alpha(V)\}.$$

Consider the constraints in  $B(f_\alpha, \geq)$ . Let

$$f_\alpha^\#(X) = f_\alpha(V) - f_\alpha(V \setminus X) = \alpha - f_\alpha(V \setminus X), \forall X \subseteq V$$

be the *dual set function* of  $f_\alpha$ . If we restrict the rate vector  $r_V$  to satisfy  $r(X) \geq f_\alpha(X)$  for some  $X \subseteq V$  and the sum-rate  $r(V) = \alpha$ , then we necessarily put constraint

$$r(V \setminus X) = r(V) - r(X) \leq f_\alpha^\#(V \setminus X)$$

on set  $V \setminus X$ . By converting the constraints in  $B(f_\alpha, \geq)$  in the same way for all  $X \subseteq V$ , we get the base polyhedron

$$B(f_\alpha^\#, \leq) = \{r_V \in \mathbb{R}_+^{|V|} : r_V \in P(f_\alpha^\#, \leq), r(V) = f_\alpha^\#(V)\}$$

such that  $B(f_\alpha^\#, \leq) = B(f_\alpha, \geq)$ .

**Example III.1.** For the CO system in Example II.1, we have  $\mathcal{C}_G = B(f_\alpha, \geq)$  where

$$\begin{aligned} B(f_\alpha, \geq) &= \{r_V \in \mathbb{R}_+^{|V|} : r(\emptyset) = 0, r(\{1\}) \geq 1, \\ &\quad r(\{2\}) \geq 0, r(\{3\}) \geq 0, \\ &\quad r(\{1, 2\}) \geq 3, r(\{1, 3\}) \geq 3 \\ &\quad r(\{2, 3\}) \geq 1, r(\{1, 2, 3\}) = \alpha\}. \end{aligned}$$

By converting constraint  $r(X) \geq f_\alpha(X)$  to  $r(V \setminus X) \leq f_\alpha^\#(V \setminus X)$  for all  $X \subseteq V$ , we have

$$\begin{aligned} B(f_\alpha^\#, \leq) &= \{r_V \in \mathbb{R}_+^{|V|} : r(\emptyset) = 0, r(\{1\}) \leq \alpha - 1, \\ &\quad r(\{2\}) \leq \alpha - 3, r(\{3\}) \leq \alpha - 3, \\ &\quad r(\{1, 2\}) \leq \alpha, r(\{1, 3\}) \leq \alpha \\ &\quad r(\{2, 3\}) \leq \alpha - 1, r(\{1, 2, 3\}) = \alpha\}. \end{aligned}$$

It can be shown that  $B(f_\alpha^\#, \leq) = B(f_\alpha, \geq), \forall \alpha \in \mathbb{R}_+$ .

### B. Interpretation of Core in CO

The core as expressed in (3) is in fact the set that contains all achievable rate vectors having sum-rate equal to  $\alpha$  in CO. It can be interpreted as follows. Since  $\mathcal{C}_G = B(f_\alpha^\#, \leq)$ , we can write  $\mathcal{C}_G$  as

$$\mathcal{C}_G = \{r_V \in \mathbb{R}_+^{|V|} : r(X) \leq f_\alpha(V) - f_\alpha(V \setminus X), \forall X \subseteq V, r(V) = \alpha\}.$$

If  $f_\alpha(X)$  is the payoff for forming coalition  $X$ , then  $f_\alpha(V) - f_\alpha(V \setminus X)$  is the cost when coalition  $X$  choose not to cooperate in the grand coalition  $V$  [13]. A rate vector  $r_V \in \mathbb{R}_+^{|V|}$  is a rate allocation method that divides sum-rate  $r(V) = \alpha$  in  $V$ , where  $r(X)$ , the sum-rate in coalition  $X$ , can be considered as the cost for  $X$  to be cooperative in  $V$ . Then,  $r(X) \leq f_\alpha(V) - f_\alpha(V \setminus X)$  means that the cost when  $X$  agrees to cooperate in  $V$  is no greater than when  $X$  denies to do so. If the core is nonempty, there exists at least one rate allocation method such that all users would not prefer forming the coalitions smaller than the grand one, or, alternatively speaking, there exists a rate allocation method that motivates all users to participate in information exchanging for CO. In this case, the game is called *stable* [14] and the core can be considered as the solution set for the game.

## IV. NONEMPTINESS OF CORE

Since the core is not guaranteed to be nonempty in all coalitional games, there is a fundamental question:

(a) When is the core nonempty?

If the core is nonempty, we need to answer the question:

(b) Can we find a rate vector  $r_V$  in the core that allocates the value of the grand coalition  $f_\alpha(V) = \alpha$  fairly among the users?

The main purpose of this section is to answer question

(a). We study the submodularity of the base polyhedron of the characteristic function to derive a necessary and sufficient condition for the nonemptiness of the core. Question (b) will be answered in Section V.

### A. Necessary and Sufficient Condition

Recall that  $\mathcal{C}_G = B(f_\alpha, \geq) = B(f_\alpha^\#, \leq)$ , i.e., we can study either  $B(f_\alpha, \geq)$  or  $B(f_\alpha^\#, \leq)$  in order to determine the nonemptiness of  $\mathcal{C}_G$ . In this section, we choose to consider  $B(f_\alpha^\#, \leq)$  based on which we show that the condition for the nonemptiness of the core can be straightforwardly derived.

**Definition IV.1** (sub/supermodular). *Function  $g: 2^V \mapsto \mathbb{R}$  is submodular if*

$$g(X) + g(Y) \geq g(X \cup Y) + g(X \cap Y), \quad \forall X, Y \subseteq V; \quad (4)$$

*$g$  is supermodular if  $-g$  is submodular.*

**Lemma IV.2.** *If  $\alpha \geq H(Z_V)$ ,  $f_\alpha^\#$  is submodular, that is*

$$f_\alpha^\#(X) + f_\alpha^\#(Y) \geq f_\alpha^\#(X \cup Y) + f_\alpha^\#(X \cap Y) \quad (5)$$

*for all  $X, Y \subseteq V$ ; If  $\alpha < H(Z_V)$ ,  $f_\alpha^\#$  is intersecting submodular, that is inequality (5) holds for all  $X, Y \subseteq V$  such that  $X \cap Y \neq \emptyset$ .*

*Proof:* For function  $f_\alpha^\#$ , we have

$$f_\alpha^\#(X) + f_\alpha^\#(Y) - f_\alpha^\#(X \cup Y) - f_\alpha^\#(X \cap Y) = \begin{cases} H(X) + H(Y) \\ -H(X \cup Y) - H(X \cap Y) + \alpha - H(V) & X \cap Y = \emptyset \\ H(X) + H(Y) - H(X \cup Y) - H(X \cap Y) & \text{otherwise} \end{cases}$$

Due to the submodularity of the entropy function  $H$ , i.e.,

$$H(X) + H(Y) \geq H(X \cup Y) + H(X \cap Y) \quad \forall X, Y \subseteq V,$$

if  $\alpha \geq H(Z_V)$ , inequality (5) holds  $\forall X, Y \subseteq V$ , i.e.,  $f_\alpha^\#$  is submodular; if  $\alpha < H(Z_V)$ , inequality (5) holds  $\forall X, Y \subseteq V: X \cap Y \neq \emptyset$ , i.e.,  $f_\alpha^\#$  is intersecting submodular. ■

Denote  $\Pi(V)$  the partition set that contains all possible partitions of  $V$  and  $\Pi'(V) = \Pi(V) \setminus \{V\}$ .

**Theorem IV.3.** *The core of  $G(V, \alpha, f_\alpha)$  is nonempty if and only if*

$$\alpha = \min_{\mathcal{P} \in \Pi(V)} \sum_{C \in \mathcal{P}} f_\alpha^\#(C). \quad (6)$$

*Proof:* Recall that  $\mathcal{C}_G = B(f_\alpha^\#, \leq)$ . If  $\alpha \geq H(Z_V)$ ,  $f_\alpha^\#$  is submodular and  $f_\alpha^\#(V) = \min_{\mathcal{P} \in \Pi(V)} \sum_{C \in \mathcal{P}} f_\alpha^\#(C) = \alpha$ . Then,  $B(f_\alpha^\#, \leq)$  is a submodular base polyhedron which is not empty [10]. If  $\alpha < H(Z_V)$ ,  $f_\alpha^\#$  is intersecting submodular, and (6) is the necessary and sufficient condition for  $B(f_\alpha^\#, \leq)$  to be nonempty according to Lemma A.1 in Appendix A. Therefore, theorem holds. ■

### B. Minimum sum-rate in CO

Based on Theorem (IV.3), we can derive the minimum sum-rate for CO as follows.

**Corollary IV.4.** *The core of  $G(V, \alpha, f_\alpha)$  is non-empty if  $\alpha \geq R_{CO}(Z_V)$ , where*

$$R_{CO}(Z_V) = \max_{\mathcal{P} \in \Pi'(V)} \sum_{C \in \mathcal{P}} \frac{H(Z_{V \setminus C} | Z_C)}{|\mathcal{P}| - 1} \quad (7)$$

*for asymptotic model and*

$$R_{CO}(Z_V) = \left\lceil \max_{\mathcal{P} \in \Pi'(V)} \sum_{C \in \mathcal{P}} \frac{H(Z_{V \setminus C} | Z_C)}{|\mathcal{P}| - 1} \right\rceil \quad (8)$$

*for non-asymptotic model.*

*Proof:* Since  $f_\alpha^\#(V) = \alpha$ , (6) in Theorem IV.3 is equivalent to  $\alpha \leq \min_{\mathcal{P} \in \Pi'(V)} \sum_{C \in \mathcal{P}} f_\alpha^\#(C)$  which can be written as

$$\alpha \geq \max_{\mathcal{P} \in \Pi'(V)} \sum_{C \in \mathcal{P}} \frac{H(Z_{V \setminus C} | Z_C)}{|\mathcal{P}| - 1}. \quad (9)$$

Then, the minimum sum-rate must be the minimum value of  $\alpha$  such that (9) holds. So, we have (7) for asymptotic model. For non-asymptotic setting,  $\alpha$  is the least integer number such that (9) holds. So, we have (8). According to Theorem IV.3,  $\mathcal{C}_G \neq \emptyset$  if  $\alpha \geq R_{CO}(Z_V)$ . ■

**Example IV.5.** *For the CO model in Example II.1, it can be show that  $R_{CO}(Z_V) = 3.5$  and  $R_{CO}(Z_V) = 4$  for asymptotic and non-asymptotic models, respectively, by applying (7) and*

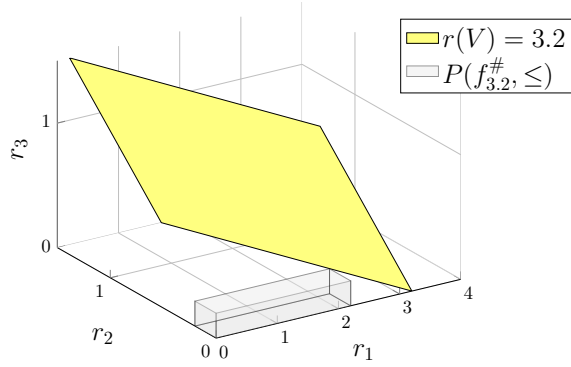


Fig. 1. The polyhedron  $P(f_{3.2}^\#, \le)$  for the CO system in Example II.1 and the plane  $r(V) = 3.2$ . In this case, the intersection  $\mathcal{C}_G = B(f_{3.2}^\#, \le) = \emptyset$ .

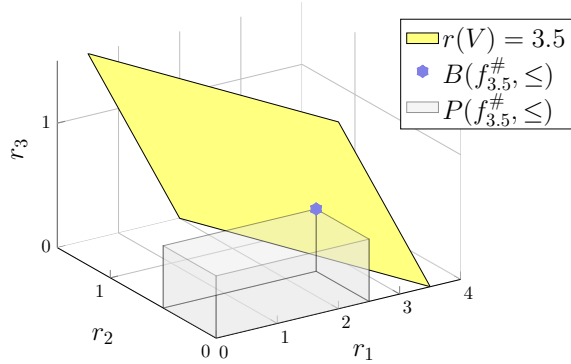


Fig. 2.  $B(f_{3.5}^\#, \le) = \{(2.5, 0.5, 0.5)\}$  for the CO system in Example II.1. In this system,  $R_{\text{CO}}(Z_V) = 3.5$  for asymptotic model. Consider game  $G(V, 3.2, f_{3.2})$ . We have  $\mathcal{R}_{\text{CO}}^*(Z_V) = \mathcal{C}_G = B(f_{3.5}^\#, \le)$  and  $\{(2.5, 0.5, 0.5)\}$  is the only one optimal rate vector for CO.

(8). If we increase  $\alpha$  from  $\alpha = 0$  and observe the base polyhedron  $B(f_\alpha^\#, \le)$ , the intersection of  $P(f_\alpha^\#, \le)$  and plane  $r(V) = \alpha$ , it can be shown that  $P(f_\alpha^\#, \le)$  does not intersect with plane  $r(V) = \alpha$ , i.e.,  $B(f_\alpha^\#, \le) = \emptyset$ , until  $\alpha = 3.5$ . For example,  $B(f_{3.2}^\#, \le) = \emptyset$  in Fig. 1. At  $\alpha = 3.5$ ,  $B(f_{3.5}^\#, \le) = \{(2.5, 0.5, 0.5)\}$  in Fig. 2 and  $B(f_\alpha^\#, \le) \neq \emptyset$  for all  $\alpha \geq 3.5$ . But,  $B(f_\alpha^\#, \le) \cap \mathbb{Z}^3 = \emptyset$  until  $\alpha = 4$  where  $B(f_4^\#, \le) \cap \mathbb{Z}^3 = \{(3, 0, 1), (2, 1, 1), (3, 1, 0)\}$  as shown in Fig. 3.

If we replace  $H(Z_{V \setminus C} | Z_C)$  with the cardinality function  $|W_V| - |W_C|$  in (7) and (8), we get exactly the minimum sum-rate expressions for asymptotic and non-asymptotic models, respectively, for CCDE in [4]. Let  $I(Z_V)$  be the MMI measure proposed in [7] that is defined as

$$I(Z_V) = \min_{\mathcal{P} \in \Pi(V)} \frac{D(P_{Z_V} \| \prod_{C \in \mathcal{P}} P_{Z_C})}{|\mathcal{P}| - 1}.$$

$D(\cdot \| \cdot)$  is the Kullback-Leibler divergence and  $D(P_{Z_V} \| \prod_{C \in \mathcal{P}} P_{Z_C}) = \sum_{C \in \mathcal{P}} H(Z_C) - H(Z_V)$ . We can write (7) and (8) as

$$R_{\text{CO}}(Z_V) = H(Z_V) - I(Z_V) \quad (10)$$

and

$$R_{\text{CO}}(Z_V) = H(Z_V) - \lfloor I(Z_V) \rfloor,$$

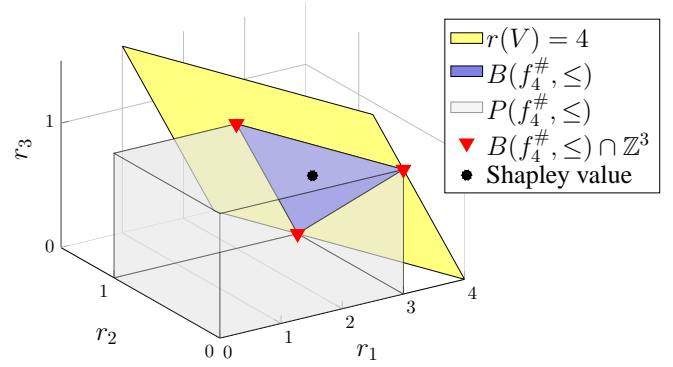


Fig. 3.  $B(f_4^\#, \le)$  for the CO system in Example II.1. In this system,  $R_{\text{CO}}(Z_V) = 4$  for non-asymptotic model. Consider game  $G(V, 4, f_4)$ . We have  $\mathcal{R}_{\text{CO}}^*(Z_V) \cap \mathbb{Z}^3 = \mathcal{C}_G \cap \mathbb{Z}^3 = B(f_4^\#, \le) \cap \mathbb{Z}^3 = \{(3, 0, 1), (2, 1, 1), (3, 1, 0)\}$ . The Shapley value calculated in Example V.3 lies in the core.

which are exactly the minimum sum-rate for CO for asymptotic and non-asymptotic models, respectively, in [6], [15]. The interpretation of (10) is as follows.  $I(Z_V)$  can be considered as the maximum amount of information that is mutual to users in  $V$  [7]. So, the minimum sum-rate for  $R_{\text{CO}}(Z_V)$  must be  $H(Z_V) - I(Z_V)$ , the amount of information that is not mutual to users in  $V$ .

### C. Convexity of Game

Convex game is a special class of coalitional game.

**Definition IV.6** (Convex Game [16]). A coalitional game is convex if the characteristic function is supermodular.

Convex game has nice properties [17]:

- The core is nonempty;
- Shapley value lies in the core.

In this section, we use the first property to interpret Corollary IV.4. The second property will be used to present a fair distribution of the value of the grand coalition  $f_\alpha(V) = \alpha$  that lies in the core in Section V.

**Lemma IV.7.** For each  $\alpha \geq R_{\text{CO}}(Z_V)$ , there exists a convex game  $\hat{G}(V, \alpha, \hat{f}_\alpha)$  such that the cores of  $G$  and  $\hat{G}$  are equal.

*Proof:* The Dilworth truncation of  $f_\alpha^\#$  is given by [18]

$$\hat{f}_\alpha^\#(X) = \min_{\mathcal{P} \in \Pi(X)} \sum_{C \in \mathcal{P}} f_\alpha^\#(C), \forall X \subseteq V. \quad (11)$$

For  $\alpha \geq H(Z_V)$ ,  $\hat{f}_\alpha^\# = f_\alpha^\#$  since  $f_\alpha^\#$  is submodular; For  $R_{\text{CO}}(Z_V) \leq \alpha < H(Z_V)$ ,  $\hat{f}_\alpha^\#$  is submodular,  $\hat{f}_\alpha^\#(V) = f_\alpha^\#(V) = \alpha$  and  $B(\hat{f}_\alpha^\#, \le) = B(f_\alpha^\#, \le)$  [10]. Let

$$\hat{f}_\alpha(X) = \hat{f}_\alpha^\#(V) - \hat{f}_\alpha^\#(V \setminus X).$$

$\hat{f}_\alpha$  is supermodular and  $B(\hat{f}_\alpha, \ge) = B(f_\alpha^\#, \le)$ . Here,  $B(\hat{f}_\alpha, \ge)$  is the core of game  $\hat{G}(V, \hat{f}_\alpha, \alpha)$ . Recall that  $\mathcal{C}_G = B(f_\alpha^\#, \le)$ . Therefore,  $\mathcal{C}_G = \mathcal{C}_{\hat{G}}$ . In addition, if  $f_\alpha^\#$  is integer-valued, so is  $\hat{f}_\alpha^\#$  [10]. Therefore,  $\mathcal{C}_G \cap \mathbb{Z}^{|V|} = \mathcal{C}_{\hat{G}} \cap \mathbb{Z}^{|V|}$  for non-asymptotic model. ■

We use Lemma IV.7 to interpret Corollary IV.4 as follows. If  $\alpha \geq R_{\text{CO}}(Z_V)$ , game  $G(V, \alpha, f_\alpha)$  is equivalent to convex game

$\hat{G}(V, \alpha, \hat{f}_\alpha)$ , where  $\hat{f}_\alpha$  can be obtained from  $f_\alpha$  by Dilworth truncation. In this case,  $\mathcal{C}_G = \mathcal{C}_{\hat{G}} \neq \emptyset$  for asymptotic model and  $\mathcal{C}_G \cap \mathbb{Z}^{|V|} = \mathcal{C}_{\hat{G}} \cap \mathbb{Z}^{|V|} \neq \emptyset$  for non-asymptotic model due to the convexity of  $\hat{G}$ .

Since function  $\hat{f}_\alpha^\#$  is submodular, the optimization problem on the core  $\mathcal{C}_G = B(\hat{f}_\alpha^\#, \leq)$  is closely related to SFM that can be solved in strongly polynomial time. For example, the problem of finding a rate vector  $r_V \in \mathcal{C}_G = B(\hat{f}_\alpha^\#, \leq)$  can be solved by Edmond's greedy algorithm [9]. Related algorithms can be found in [3], [8]. These algorithms always find an extreme point (a vertex) of the core.

**Example IV.8.** Consider the CO system in Example II.1 when  $\alpha = 4$  and its corresponding game model  $G(V, 4, f_4)$ . We obtain  $\hat{f}_4^\#$  by the Dilworth truncation (11) and obtain  $\hat{f}_4$  as

$$\begin{aligned} \hat{f}_4(\emptyset) &= 0, \hat{f}_4(\{1\}) = 2, \hat{f}_4(\{2\}) = 0, \hat{f}_4(\{3\}) = 0, \\ \hat{f}_4(\{1, 2\}) &= 3, \hat{f}_4(\{1, 3\}) = 3, \hat{f}_4(\{2, 3\}) = 1, \\ \hat{f}_4(\{1, 2, 3\}) &= 4. \end{aligned} \quad (12)$$

$\hat{f}_4$  is supermodular and therefore  $\hat{G}(V, 4, \hat{f}_4)$  is a convex game. It can be shown that  $\mathcal{C}_G = B(f_4, \geq) = B(\hat{f}_4, \geq) = \mathcal{C}_{\hat{G}}$  and  $\mathcal{C}_G \cap \mathbb{Z}^{|V|} = B(f_4, \geq) \cap \mathbb{Z}^{|V|} = B(\hat{f}_4, \geq) \cap \mathbb{Z}^{|V|} = \mathcal{C}_{\hat{G}} \cap \mathbb{Z}^{|V|}$ . In addition,  $B(\hat{f}_4, \geq) = B(\hat{f}_4^\#, \leq) = B(f_4^\#, \leq)$ , where  $B(f_4^\#, \leq)$  is shown in Fig. 3.

## V. SHAPLEY VALUE

To make sure that all users are willing to cooperate, it is not sufficient to just know that the core is nonempty. Since the core is not a singleton in general, some solutions in the core may not be a good choice in terms of fairness. For example, for the CO system in Example II.1 when  $\alpha = 4$ , the Edmond's greedy algorithm usually returns a vertex of the core  $\mathcal{C}_G = B(f_4^\#, \leq)$ , e.g.,  $(3, 0, 1)$ , which is one of the unfairest solutions in  $\mathcal{C}_G$  that may make one or more users unwilling to cooperate. In CCDE, to achieve some degree of fairness in rate allocation can also prevent running out of mobile clients' battery usage.

The authors in [19], [20] proposed polynomial time algorithms to compute a fair rate vector in core  $\mathcal{C}_G$  for non-asymptotic model in CCDE. The greedy algorithm in [19] is based on SFM. The authors in [20] have shown that a fair rate in the base polyhedron  $B(f_\alpha^\#, \leq)$ , or the core  $\mathcal{C}_G$ , can be found by solving an  $M$ -convex minimization for which there exists a discrete steepest descent algorithm that can search the optimal solution in polynomial time.

Finding a fair rate vector for the asymptotic model is more complex than non-asymptotic one. Although the problem can be easily formulated by a convex minimization problem, e.g.,  $\min\{\sum_{i \in V} r_i^2 : r_V \in \mathcal{C}_G\}$ , where the objective function is defined based on Jain's fairness index [21], the number of constraints  $2^{|V|}$  is exponentially growing in  $|V|$ . In [11], the authors build a multi-layer graph model and formulate a convex minimization problem based on it. Although the constraints is not as large as  $2^{|V|}$ , building new model and deriving constraints based on this model incurs extra complexity. In this section, we show an alternative way to achieve the fairness based on game model  $G(V, \alpha, f_\alpha)$ : Shapley value.

**Definition V.1** (Shapley Value [22]). For a coalitional game with characteristic function  $g: 2^V \mapsto \mathbb{R}$ , the Shapley Value is a rate vector  $\hat{r}_V$  with each entry being

$$\hat{r}_i = \sum_{X \subseteq V: i \in X} \frac{(|X| - 1)!|V \setminus X|!}{|V|!} (g(X) - g(X \setminus \{i\})).$$

The weight factor  $\frac{(|X| - 1)!|V \setminus X|!}{|V|!}$  is the probability for user  $i$  to enter coalition  $X$  in a random order [22]. Therefore,  $\hat{r}_i$  can be considered as the expected marginal value of the characteristic function  $g$  when the users join coalitions randomly to form the grand coalition  $V$  [14].

**Lemma V.2.** For asymptotic model, if  $\alpha \geq R_{CO}(Z_V)$ , the Shapley value  $\hat{r}_V$  with

$$\hat{r}_i = \sum_{X \subseteq V \setminus \{i\}} \frac{(|V \setminus X| - 1)!|X|!}{|V|!} (\hat{f}_\alpha^\#(X \cup \{i\}) - \hat{f}_\alpha^\#(X))$$

lies in the core and is a fair method to distribute the value of the grand coalition  $f_\alpha(V) = \alpha$  among all users.

*Proof:* Based on the proof of Lemma IV.7, if  $\alpha \geq R_{CO}(Z_V)$ ,  $\mathcal{C}_G = \mathcal{C}_{\hat{G}}$ . Game  $\hat{G}(V, \alpha, \hat{f}_\alpha)$  is convex and the characteristic function  $\hat{f}_\alpha$  is defined as  $\hat{f}_\alpha(X) = \hat{f}_\alpha^\#(V) - \hat{f}_\alpha^\#(V \setminus X)$ ,  $\forall X \subseteq V$ . The Shapley value of  $\hat{G}$  is

$$\begin{aligned} \hat{r}_i &= \sum_{X \subseteq V: i \in X} \frac{(|X| - 1)!|V \setminus X|!}{|V|!} (\hat{f}_\alpha(X) - \hat{f}_\alpha(X \setminus \{i\})) \\ &= \sum_{X \subseteq V \setminus \{i\}} \frac{(|V \setminus X| - 1)!|X|!}{|V|!} (\hat{f}_\alpha^\#(X \cup \{i\}) - \hat{f}_\alpha^\#(X)). \end{aligned}$$

It is shown in [23] that  $\hat{r}_V \in \mathcal{C}_{\hat{G}}$  and is a fair method to distribute  $\alpha$  in set  $V$ . ■

**Example V.3.** Consider the CO system in Example II.1. We obtain function  $\hat{f}_4^\#$  by Dilworth truncation (11) as

$$\begin{aligned} \hat{f}_4^\#(\emptyset) &= 0, \hat{f}_4^\#(\{1\}) = 3, \hat{f}_4^\#(\{2\}) = 1, \hat{f}_4^\#(\{3\}) = 1, \\ \hat{f}_4^\#(\{1, 2\}) &= 4, \hat{f}_4^\#(\{1, 3\}) = 4, \hat{f}_4^\#(\{2, 3\}) = 2, \\ \hat{f}_4^\#(\{1, 2, 3\}) &= 4. \end{aligned}$$

Note, the values of  $\hat{f}_4^\#$  have been used to determine function  $\hat{f}_4$  in Example IV.8. For user 1, the possible values of  $X \subseteq V \setminus \{1\}$  are  $\emptyset$ ,  $\{2\}$ ,  $\{3\}$  and  $\{2, 3\}$ . Therefore,

$$\begin{aligned} \hat{r}_1 &= \frac{2!0!}{3!}(3 - 0) + \frac{1!1!}{3!}(4 - 1) \\ &\quad + \frac{1!1!}{3!}(4 - 1) + \frac{0!2!}{3!}(4 - 2) = \frac{8}{3}. \end{aligned} \quad (13)$$

In the same way, we find  $\hat{r}_2 = \frac{2}{3}$  and  $\hat{r}_3 = \frac{2}{3}$ . The Shapley value  $\hat{r}_V = (\frac{8}{3}, \frac{2}{3}, \frac{2}{3})$  is plotted in Fig. 3, where it can be seen that  $\hat{r}_V \in \mathcal{C}_G$  and  $\hat{r}_V$  is a fair rate allocation method as compared to other rate vectors in the core.

Although obtaining the Dilworth truncation  $\hat{f}^\#(X)$  for a particular  $X$  is related to SFM and can be completed in strongly polynomial time [24], the complexity of obtaining Shapley value is exponentially growing with  $|V|$ . How to reduce the complexity of calculating Shapley value is still an open problem in the literature [22]. However, the advantage

of Shapley value over the method in [11] is that it can be learned in a decentralized manner: As long as user  $i$  obtains the knowledge of the whole system, e.g.,  $f_\alpha$ , his/her tuple  $\hat{r}_i$  in Shapley value can be calculated by him/herself.

## VI. CONCLUSION

We formulated the problem of CO by a coalition game model where the core contained all achievable rate vectors that satisfied the Slepian-Wolf constraints for CO and had sum-rate equal to a given value. We showed that the core was a base polyhedron of a submodular or intersecting submodular function. We derived the necessary and sufficient condition for the core to be nonempty, based on which we gave the expression of the minimum sum-rate for CO and showed that they were consistent with the existing results in the literature. We proved that the game was convex when the sum-rate was greater or equal to the minimum sum-rate and showed that a fair rate vector in the core could be obtained by calculating the Shapley value. However, since the complexity of obtaining the Shapley value is exponentially growing, it is worth discussing how to allow users to learn the fair rate vector in the core in polynomial time, which could be one of the direction of research work in the future.

## APPENDIX A

**Lemma A.1.** *For an intersecting submodular function  $g: 2^V \mapsto \mathbb{R}_+$  such that  $g(\emptyset) = 0$ , the base polyhedron  $B(g, \leq)$  is nonempty if and only if*

$$g(V) = \min_{P \in \Pi(V)} \sum_{C \in P} g(C)$$

*Proof:* Theorem 2.6 in [10] gives the necessary and sufficient conditions for  $B(g, \leq)$  to be nonempty:

$$g(V) = \max_{P \in \Pi(V)} \sum_{C \in P} g^\#(C) = \min_{P \in \Pi(V)} \sum_{C \in P} g(C), \quad (14)$$

which is equivalent to  $g(V) \geq \sum_{C \in P} g^\#(C)$  and  $g(V) \leq \sum_{C \in P} g(C)$  for all  $P \in \Pi'(V)$ . The latter can be written as

$$\begin{aligned} g(V) &\leq \sum_{C \in P} (g(V) - g^\#(V \setminus C)) \\ \Rightarrow g(V) &\leq |P|g(V) - \sum_{C \in P} g^\#(V \setminus C) \\ \Rightarrow g(V) &\geq \frac{\sum_{C \in P} g^\#(V \setminus C)}{|P| - 1}, \quad \forall P \in \Pi'(V). \end{aligned}$$

Due to the intersecting submodularity of  $g$ , for all  $\emptyset \neq X, Y \subset V$  such that  $X \cap Y = \emptyset$  and  $X \cup Y \neq V$ , we have

$$g^\#(X \cup Y) \geq g^\#(X) + g^\#(Y).$$

So, for all  $C \in P$  where  $P$  is some partition in  $\Pi'(V)$ ,  $g^\#(V \setminus C) \geq \sum_{C' \in P \setminus C} g^\#(C')$  and  $\sum_{C \in P} g^\#(V \setminus C) \geq (|P| - 1) \sum_{C \in P} g^\#(C)$ , i.e.,

$$\sum_{C \in P} \frac{g^\#(V \setminus C)}{|P| - 1} \geq \sum_{C \in P} g^\#(C), \quad \forall P \in \Pi'(V).$$

Therefore, (14) reduces to  $g(V) \leq \sum_{C \in P} g(C), \forall P \in \Pi'(V)$ , which is equivalent to  $g(V) = \min_{P \in \Pi(V)} \sum_{C \in P} g(C)$ . ■

## REFERENCES

- [1] I. Csiszár and P. Narayan, "Secrecy capacities for multiple terminals," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3047–3061, Dec. 2004.
- [2] S. El Rouayheb, A. Sprintson, and P. Sadeghi, "On coding for cooperative data exchange," in *Proc. IEEE Inform. Theory Workshop*, Cairo, 2010, pp. 1–5.
- [3] T. Courtade and R. Wesel, "Coded cooperative data exchange in multi-hop networks," *IEEE Trans. Inf. Theory*, vol. 60, no. 2, pp. 1136–1158, Feb. 2014.
- [4] N. Ding, R. A. Kennedy, and P. Sadeghi, "Estimating minimum sum-rate for cooperative data exchange," in *Proc. IEEE Int. Symp. Inform. Theory*, Hong Kong, China, 2015. [Online]. Available: <http://arxiv.org/abs/1502.03518>
- [5] C. Chan, "Linear perfect secret key agreement," in *Proc. IEEE Inform. Theory Workshop*, Paraty, Oct 2011, pp. 723–726.
- [6] C. Chan, Al-Bashabsheh, Q. Zhou, N. Ding, T. Liu, and A. Sprintson, "Successive omniscience," *submitted to IEEE Trans. Inf. Theory*, 2015.
- [7] C. Chan, A. Al-Bashabsheh, J. Ebrahimi, T. Kaced, and T. Liu, "Multivariate mutual information inspired by secret-key agreement," *Proc. of the IEEE*, vol. 103, no. 10, pp. 1883–1913, Oct. 2015.
- [8] N. Milosavljevic, S. Pawar, S. E. Rouayheb, M. Gastpar, and K. Ramchandran, "Efficient algorithms for the data exchange problem," *arXiv preprint arXiv:1502.06149*, 2015.
- [9] J. Edmonds, "Submodular functions, matroids, and certain polyhedra," in *Combinatorial Optimization—Eureka, You Shrink!* Berlin: Springer, 2003, pp. 11–26.
- [10] S. Fujishige, *Submodular functions and optimization*, 2nd ed. Amsterdam, The Netherlands: Elsevier, 2005.
- [11] S. Tajbakhsh, P. Sadeghi, and R. Shams, "A generalized model for cost and fairness analysis in coded cooperative data exchange," in *Proc. Int. Symp. Network Coding*, Beijing, 2011, pp. 1–6.
- [12] R. W. Yeung, *Information theory and network coding*. Berlin, Germany: Springer Science & Business Media, 2008.
- [13] J. M. Bilbao, *Cooperative games on combinatorial structures*. Springer Science & Business Media, 2012, vol. 26.
- [14] L. S. Shapley and M. Shubik, "On market games," *J. Economic Theory*, vol. 1, no. 1, pp. 9–25, Jun. 1969.
- [15] C. Chan, A. Al-Bashabsheh, J. B. Ebrahimi, T. Kaced, S. Kadhe, T. Liu, A. Sprintson, M. Yan, and Q. Zhou, "Successive omniscience," in *Proc. Int. Symp. Network Coding (NetCod)*, Sydney, 2015, pp. 21–25.
- [16] Y. Shoham and K. Leyton-Brown, *Multiagent systems: Algorithmic, game-theoretic, and logical foundations*. New York: Cambridge University Press, 2008.
- [17] L. S. Shapley, "Cores of convex games," *Int. J. Game Theory*, vol. 1, no. 1, pp. 11–26, Dec. 1971.
- [18] R. P. Dilworth, "Dependence relations in a semi-modular lattice," *Duke Math. J.*, vol. 11, no. 3, pp. 575–587, 1944.
- [19] N. Milosavljevic, S. Pawar, M. Gastpar, and K. Ramchandran, "Efficient algorithms for the data exchange problem under fairness constraints," in *Proc. Annu. Allerton Conf. Commun., Control, and Comput.*, Monticello, IL, 2012, pp. 502–508.
- [20] N. Ding, R. A. Kennedy, and P. Sadeghi, "Fairness constant sum-rate transmission for cooperative data exchange: An  $M$ -convex minimization approach," in *Proc. 22nd Int. Conf. Telecommun.*, Sydney, Australia, 2015. [Online]. Available: <http://arxiv.org/abs/1502.03517>
- [21] R. Jain, D.-M. Chiu, and W. R. Hawe, *A quantitative measure of fairness and discrimination for resource allocation in shared computer system*. Hudson, MA: Eastern Research Laboratory, Digital Equipment Corporation, 1984.
- [22] W. Saad, Z. Han, M. Debbah, A. Hjørungnes, and T. Başar, "Coalitional game theory for communication networks," *IEEE Signal Process. Mag.*, vol. 26, no. 5, pp. 77–97, 2009.
- [23] T. Lan, D. Kao, M. Chiang, and A. Sabharwal, "An axiomatic theory of fairness in network resource allocation," in *Proc. IEEE INFOCOM*, San Diego, CA, 2010, pp. 1–9.
- [24] A. Schrijver, *Combinatorial Optimization: Polyhedra and Efficiency*. Berlin: Springer-Verlag, 2003.